A discusión

COMMUTING COSTS AND LABOR FORCE RETIREMENT*

Jorge González

WP-AD 2008-19

Correspondence: Universidad de Alicante. Fundamentos del Análisis Económico. Campus San Vicente del Raspeig, s/n. 03071 Alicante. Tel.: +34 965 90 3400 x 3263/ Fax: +34 96 5903898 / E-mail: jorge@merlin.fae.ua.es.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
Primera Edición Diciembre 2008
Depósito Legal: V-5153-2008

* I am grateful to Daniel Hamermesh for advice and encouragement at the beginning of this project; to Namkee Ahn, Axel Boersch-Supan, Martin Browning, José Canals-Cerdá, Elsa Fornero, Sergi Jiménez, José María Labeaga, Juan Mora, Pascal Mossay, Arthur van Soest, Nikos Varsakelis, Frederic Vermeulen, seminar participants at the Universidad de Alicante, the Fundación de Estudios de Economía Aplicada (FEDEA), the XXXII Simposio de Análisis Económico held in Granada, the Tilburg University, and the Work, Pensions and Labour Economics Study Group (WPEG) Conference held in Sheffield for helpful comments or suggestions; and to the Instituto Valenciano de Investigaciones Economicas (IVIE) and the Spanish Ministry of Education (SEJ2005-02829/ECON) for financial support. My deepest gratitude to Lola Collado for an effort above any call of duty. Comments gratefully received at jorge@merlin.fae.ua.es

COMMUTING COSTS AND LABOR FORCE RETIREMENT

Jorge González

ABSTRACT

This paper studies whether the increase in home-workplace separation observed among U.S. older male workers in the last decades of the 20th century can partly account for earlier retirement. We first extend a conventional residential location-labor supply model in order to examine potential mechanisms linking commuting and retirement. After showing that, as a consequence of the urban residential equilibrium, it is possible that workers residing further from the workplace retire earlier, PSID data and an instrumental variables approach are combined in order to assess the nature and strength of the relation.

Keywords: Retirement. Commuting. Instrumental variables.

JEL classification: J26, R22.

1. INTRODUCTION

The choice of the moment of retirement from the labor force is an important decision in the sphere of individual labor supply behavior, whose study can serve to validate economic theories, to know the magnitude of one facet of the labor-leisure trade-off, or to know the consequences of current public policies and inform their eventual reform. For reasons like these, economists have devoted substantial efforts towards understanding the retirement decision. A general conclusion emerging from the academic literature on retirement is that its timing is influenced by incentives stemming from many different factors, including the provisions of pension and social security systems (e.g., Boskin 1977, Mitchell and Fields 1984, Stock and Wise 1990, Gruber and Wise 1999, 2004), health conditions (Rust and Phelan 1997, French 2005), the availability of health insurance (Madrian et al. 1994), couples' synchronous enjoyment of recreation activities (Gustman and Steinmeier 2000, Coile 2004), and job characteristics (Filer and Petri 1988). Despite this seemingly ample research, a scholar in the field of aging has recently concluded that the decline in labor force participation rates (LFPR) observed among U.S. older men from the 1960s through the 1980s cannot be explained by any of the main factors investigated in that literature (Blau and Goodstein 2007).

When the residential and job locations do not coincide, and it is thus necessary to commute between them, several costs are incurred. We spend money on gas or fares, we spend time, we suffer the discomfort created by crowds, traffic congestion, weather conditions, etc. In 2003, for example, weekly commuting expenses of an U.S. worker in the age range 55-64 averaged $69,[1] whereas average commuting time for the same person was about 3 hours per week[2]. A study comparing workers' affective evaluations of 16 common daily activities (Kahneman et al. 2004) reports that commuting has the lowest rating in the scale of positive affect descriptors, and the third highest in the scale of negative affect descriptors, being for example above housework in negative affect. Intuitively, one may wonder whether the disincentives to remain employed created by

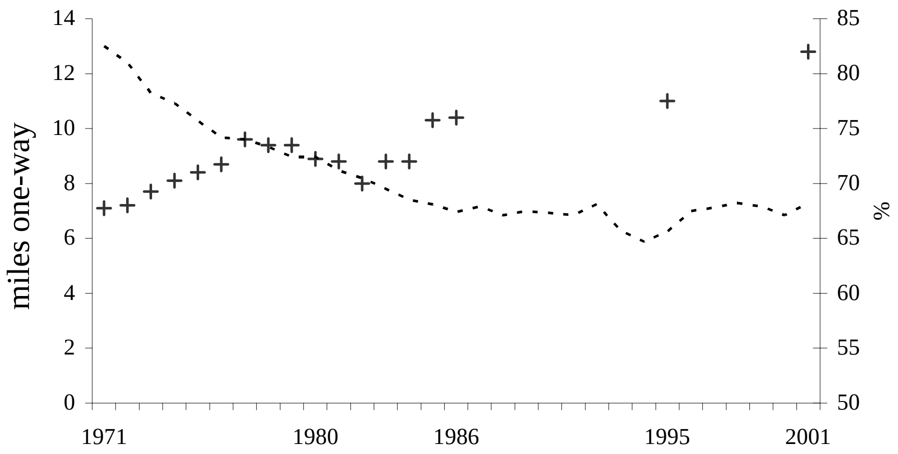
[1] Author's calculation with data from the Survey of Income and Program Participation (http://www.sipp.census.gov/sipp/, accessed on March 23, 2007). In particular, data from the 9th wave of the 2001 SIPP Panel, collected between October 2003 and January 2004, were used. Workers' commuting expenses include fares for those who did not use their own vehicle to commute, and mileage expenses plus parking fees and tolls for those who used their own vehicle. Mileage expenses are calculated as miles driven times 51.7 cents per mile. The average cost per mile is obtained from a publication of the American Automobile Association.
(http://www.aaanewsroom.net/Main/Default.asp?CategoryID=4&ArticleID=196, accessed on March 23, 2007).

[2] Author's calculation with data from the U.S. Bureau of Labor Statistics' American Time Use Survey 2003 (http://www.bls.gov/tus/, accessed on October 26, 2006).

these costs reduce the age at retirement. Indeed, as shown in Figure 1, which represents the average (one-way) distance to the workplace for the population of U.S. workers aged 55-64 for selected years between 1971 and 2001[3], and the LFPR of men 55-64, the decline in older males' LFPR in the U.S. coincided with their tendency to live further from the workplace.[4] We want to know whether this increase in home-workplace separation can partly account for earlier retirement.

Figure 1. Average home-workplace separation (+) and LFPR (-)

U.S. male workers 55-64



Notes: Author's calculations with data from: (+) Panel Study of Income Dynamics (1971-1986), Nationwide Personal Transportation Survey (1995), and National Household Travel Survey (2001); (-) Bureau of Labor Statistics.

The purpose of this paper is therefore to investigate the effect of commuting on retirement. Section 2 examines this potential link from a theoretical viewpoint, extending a residential location model with a labor-leisure choice. There we shall see that the costs of commuting are partly chosen by the worker, and that, in the urban space, these costs are compensated by housing savings. But if the costs of commuting are compensated, how can commuting affect retirement? We discuss the conditions for

[3] Values for 1971-1986 were calculated with data from the PSID (in 1969 and 1970 the PSID asked distance to the workplace of those driving to get to work only), whereas those for 1995 and 2001 were computed with data from the U.S. Department of Transportation's Nationwide Personal Transportation Survey 1995 (NPTS) and National Household Travel Survey 2001 (NHTS). NPTS and NHTS data were downloaded on August 2, 2007, starting at http://www.fhwa.dot.gov/policy/ohpi/nhts/index.htm. The NPTS 1990 does not contain information about distance to work. We think estimates are comparable, for the three surveys are representative of the non-institutionalized population and the questions used to assess the home-workplace separation do not seem capable of inducing different reporting errors: the survey question used by the PSID is "About how many miles is it to where you work?" That used by the U.S. Department of Transportation surveys is "What is the one-way distance from your home to your workplace?"

[4] This tendency is not exclusive of U.S. older workers, but part of a general shift common to OECD countries (Schafer 2000).

the existence of an age-at-retirement gradient in the urban space: motivated solely by the existence of commuting costs, workers who locate further from the workplace, but that are identical in all other respects to workers residing closer, may retire at a different age. Section 3 uses data from the Michigan Panel Study of Income Dynamics in order to assess the sign and magnitude of the effect of commuting on retirement: *ceteris paribus*, how would the age at retirement change if distance to the workplace is increased by one mile? To answer this question we make use of an instrumental variable approach, which allows overcoming the endogeneity issue created by commuting costs being in part chosen by the workers. A summary of the analysis is provided in Section 4.

This paper contributes to a field of research formed by the intersection of urban and labor economics, which is summarized and critically assessed in Simpson and van der Veen (1992), Koslowsky et al. (1995, Ch. 5), and Crampton (1999). Of the works in this field, Oi (1976) and Kolodziejczyk (2006) are particularly relevant to our research. Oi develops a theoretical framework in which workers choose residential locations and supplies of daily hours of work and workdays per year in order to maximize utility. Kolodziejczyk links the age at retirement to the fixed costs of work. He shows that, under certain conditions, higher fixed costs of work may lead to earlier retirement, particularly when workers are impatient. In his model, however, the fixed costs of work are exogenous to the worker, and there is no attempt to quantify the effect. Both concerns will be considered in our contribution.

2. A MODEL OF RESIDENTIAL LOCATION AND LABOR FORCE RETIREMENT

Considered here is the decision of a worker (or of a household represented by a worker), whose job is secured, but whose residential location is to be chosen. As in Alonso (1964), the place where the worker lives (or arrives) is conceived as a circular entity located on a uniform, featureless plain, with all employment concentrated in a central business district (CBD) of fixed and negligible size. (The centralized employment assumption is justified below.) The worker chooses one residential location, represented by its distance to the CBD. The transportation system is assumed to be infinitely-elastic supplied in any direction from the CBD, being the only travel that of workers commuting between residences and work places.

The worker's remaining years of life (T) are split into years working (H) and years being retired (L).[5] Preferences are defined over L, housing space (Q), other consumption goods (X), and distance to the CBD (D), and are represented by the utility function

$$U = U(L, Q, X, D). \tag{1}$$

Utility is assumed increasing in the first three arguments $(U_j > 0, j = 1,2,3)$, but decreasing in the fourth $(U_4 < 0)$: other things equal, the worker would prefer to live near the CBD to reduce the discomfort experienced during the commute.

The continuously differentiable function $R(D)$ stands for the market price of housing space, which the worker takes as given.[6] An extra year of work implies incurring in commuting expenses of amount $C(D)$, which are assumed increasing in D.[7] The worker's budget constraint can thus be written as

$$X + R(D)Q + C(D)H \leq WH + Y, \tag{2}$$

where W represents yearly earnings[8] and Y is non-labor income. Since $H = T - L$,

$$X + R(D)Q + (W - C(D))L \leq F \equiv (W - C(D))T + Y, \tag{3}$$

where F stands for the potential income available if working all T periods. In this model, the (opportunity) price of an extra year as retiree equals foregone earnings minus avoided commuting expenses. That the price of leisure is affected by D makes the residential location and labor supply decisions to be simultaneously chosen.

---

[5] For expository convenience, retirement is here modeled as an abrupt transition from work to complete labor-force withdrawal. Although the process by which older workers withdraw the labor market may be more complex, Lumsdaine and Mitchell (1999) concluded that the conventional pattern of work followed by non-work was characteristic of most older workers' paths out of the labor market.

[6] In a competitive equilibrium, $R(D)$ would coincide with one of the worker's price offer curves, these describing how much a worker is willing to pay for one unit of housing space at each location while enjoying a fixed level of utility (Straszheim 1987, p. 730).

[7] To concentrate the analysis, the within-period allocation of time is not specified, and the time cost of the journey to work is neglected (unless the number of commuting trips per year and the time per trip were constant, for then the time cost would be implicitly included in C). See, e.g., Oi (1976), Wales (1978), and Cogan (1981) for analyses of within-year time allocation and commuting costs.

[8] As discussed in Rosen (1986, p. 674), in a monocentric urban model with homogeneous labor a single wage clears the labor market. Our analysis further abstracts equalizing or compensating wage differentials. These differences are best interpreted as the result of a job search process in which the acceptance wage increases in commuting expenses (Rosen 1986 and Crampton 1999). In this contribution we abstract job search from the model.

In the U.S., urban employment has been suburbanizing for a long time, and the assumption of centralized employment may be controversial. It has been argued that one of the basic factors that cause employment to move out of the CBD is the firms' incentive to suburbanize as a consequence of paying lower wages, which workers are willing to accept because they would commute less. Even so, population was more suburbanized than employment during the time period (the 1970s and 1980s) object of empirical study (White 1999). Also, as long as decentralized firms capture their workers' commuting savings in the form of lower wages, the workers' price of leisure and, therefore, their demand for leisure would remain unaltered. For these reasons, centralized employment, it seems to me, is the simplest assumption mathematically that embodies the essential economics of the problem.

The worker's decision-making process can be decomposed into two stages. In the first stage, optimal quantities of L, Q, and X are chosen for a given distance, obtaining a constrained or short-run maximum of utility. In the second stage, distance itself is chosen so as to maximize utility. Hence, the equilibrium of the model is characterized by expression (3), a set of first-order conditions for the short-run optimum, and a condition for the second-stage equilibrium distance:[9]

$$-QR' = -\frac{U_4}{\lambda} + HC' . \tag{4}$$

The right-hand side of (4) represents the commuting costs (respectively, commuting savings) of a very short move from the equilibrium distance and away from (towards) the CBD. Marginal commuting costs are here made up of marginal commuting expenses (HC') as well as marginal disutility costs ($-U_4/\lambda$, expressed in monetary terms using the marginal utility of income, λ). For the term on the left-hand side to be positive too, housing prices must decline with distance to the CBD, $R' < 0$, a result supported by the majority of the empirical evidence collected (Ball 1973, Sheppard 1999). Thus, $-QR'$ denotes the savings (respectively, costs) implicit in purchasing a given quantity of housing that accrue from a very short move from the equilibrium distance and away from (towards) the CBD. Condition (4) expresses that the worker is unable to increase real income by a change of distance.

The reason why expression (4) only indicates a property of the optimal distance without determining distance itself is that, if R' satisfies (4), the worker is indifferent regarding distance, for in case of locating further, say, from the CBD greater commuting costs would be compensated by greater housing savings. But if the disincentives to remain

[9] All these conditions are derived in the technical appendix. Condition (4) is further discussed in, among others, Alonso (1964), Muth (1969), Stucker (1975), and Oi (1976).

employed created by commuting costs are offset by housing savings, how might commuting and retirement be related? In the present framework, three different mechanisms might generate a relation between both. We discuss each in turn (algebraic derivations are provided in the technical appendix).

Whenever R' satisfies (4), any distance to the CBD is of equilibrium. Yet, different distances imply different prices, and the equilibrium demand for leisure of identical workers could vary across the urban space as a consequence of substitution and income effects. Let $U_{j4} = 0,\ j = 1,2,3$.[10] Then

$$\frac{\partial L}{\partial D} = -\left(QR' + HC'\right)\mu_1 + R's_{12} - C's_{11}. \tag{5}$$

Consider a worker residing further from the CBD than other colleagues. To compensate for his greater commuting costs, his housing savings have to be greater than his commuting expenses (for commuting costs include disutility costs), thus increasing his consumption possibilities. The term $-\left(QR' + HC'\right)\mu_1$ in (5), where μ_1 is the full-income derivative of L, captures this income effect. By facing lower housing prices, this worker's demand for leisure would be also modified by a cross-price effect, represented by the term $R's_{12}$ in (5), where s_{12} is the compensated cross-price derivative of L with respect to the price of Q. Finally, by incurring in greater commuting expenses, his price of leisure would be lower. The positive term $-C's_{11}$ in (5) stands for this effect, where s_{11} is the compensated own-price derivative of L. Although the sign of (5) is an empirical matter, if L is a normal good and L and Q are Slutsky complements, expression (5) would be positive: workers residing further from the CBD, but being identical in all other respects to those residing closer, would retire earlier. This independent effect of commuting on retirement would give rise to an age-at-retirement gradient in the urban space.

The age at retirement might be also modified by a variation in the amount of commuting expenses for a given distance, $C(D)$. More efficient transport means, changes in traffic congestion, or the introduction of commuting subsidies, for example, would generate this kind of variation. Keeping distance constant, a reduction, say, in commuting expenses would cause a negative income effect on age at retirement if L is a normal good. However, lower commuting expenses would increase the price of leisure, generating a positive substitution effect on age at retirement. If, as a consequence of the

[10] That the utility of an extra year as retiree, of an extra square meter of housing space, and of an extra unit of other consumption goods are not influenced by distance is in accordance with our notion of the urban space as being featureless.

variation in commuting expenses, the worker moves,[11] new effects would further complicate the prediction of the sign of $\partial L/\partial C$.

Finally, commuting and retirement may be simultaneously affected by a common third variable. For example, it has been observed that, in the U.S., higher-income families settle further from the urban centre (e.g., Glaeser et al. 2008). If the demand for leisure depends on income, a (spurious) relation between commuting and retirement might be observed as a consequence of workers' heterogeneous incomes.

Although any of the three mechanisms just described can generate an empirical relation between commuting and retirement, both the causal status of the relation and the kind of data needed for estimating its magnitude are not the same. While the first two mechanisms can give rise to genuine causal relations between commuting and retirement, the empirical relation arising out of the last mechanism is however mediated by workers' heterogeneous characteristics, and cannot be given a causal status. Regarding data needs, a cross-section of microeconomic data representative of a given urban area is sufficient to estimate the size of the equilibrium age-at-retirement gradient for that area, provided the data include information on current location of the home and the workplace as well as expected age at retirement, or current labor force status plus recent information on the home and workplace locations. If the data represents a region comprising several urban areas, the potential heterogeneity of urban age-at-retirement gradients has to be taken into account. To be able to identify the causal relation between commuting and retirement arising out of a change in commuting expenses for a given distance, however, panel data in which the event causing the change in commuting expenses could be identified would be required.

3. THE EFFECT OF COMMUTING ON RETIREMENT

We now turn to the empirical question of estimating the independent effect of commuting on retirement, in particular that arising out of the equilibrium age-at-retirement gradient, for the U.S. economy. In this order, we shall discuss the data and sample design utilized, the econometric model assumed, and the results obtained.

3.1. Data and sample design

The Panel Study of Income Dynamics (PSID) is one of the few large-scale surveys that have collected information on labor market outcomes and commuting characteristics.

[11] Stucker (1975) studies how, and provides some evidence on by how much, a transport improvement may alter the structure of a community.

Since 1968, and generally on a yearly basis, this longitudinal survey interviews a representative sample of the U.S. population collecting a wide range of information (demographic, employment, income, etc.) In most data collection waves carried out between 1969 and 1986, heads and wives of PSID family units who were employed or looking for a job were asked the distance to the workplace and the travel to work time, both referring to the last job held if the respondent was unemployed at the moment of the interview.[12] Unfortunately, the PSID collects no information regarding the home and job locations, and therefore one cannot know whether the distance to the workplace genuinely represents distance to the CBD. An important proportion of commutes, however, should be commutes to the CBD, for Gobillon et al. (2007, p. 2404) estimate that, in the 10 largest U.S. metropolitan statistical areas, the proportion of jobs located in central cities was still 57 per cent in 1980. From this survey, all male labor force participants aged 41-60 in 1969, and considered heads of PSID families, are tracked from 1969 (the first wave with commuting information available) through 2005 (at the moment this study is done, the last wave available in the PSID). In the meantime, the study subjects may retire from the labor force, drop out from the survey, or still remain participating in 2005.

The theoretical model developed previously assumed that distance to the workplace remained the same until the worker retired. The reason for making this assumption is that, even if several residential or workplace changes take place along a worker's life, the age at which the worker leaves the work force would remain a function of the commuting characteristics of the final home-workplace separation, the separation for which commuting data are generally available. Therefore, although tracked along many periods, each study subject will contribute only one observation to this study (referred hereafter as the data wave): that corresponding to the survey wave closest to the wave of retirement that shows a change of residence or job.[13] To be consistent with this criterion, study subjects who retire after 1987, but that move or change job after the

[12] "About how much time does it take you to get to work each day, door to door?"; "About how many miles is it to where you work?" Commuting time was not asked in 1982, whereas in 1969 and 1970 commuting distance was asked of those driving to get to work only. The PSID makes available the annualized hours spent traveling to and from work. The time of a one-way commute is calculated dividing annual commuting hours by the product of weeks worked per year and commuting trips per week. Trips per week are assigned on the basis of hours worked per week on the main job according to this rule: if hours worked per week were less than 5, 2 trips per week were assumed; if hours were between 6 and 10, 4 trips were assumed; if hours were between 11 and 20, 6 trips were assumed; if hours were between 21 and 35, 8 trips were assumed; if hours were between 36 and 47, 10 trips were assumed; for 48 or more hours, 12 trips were assumed.

[13] As discussed below, to avoid a potential sample selection issue dummies for the age declared in the data wave have to be included among the covariates in the regression model for age at retirement. This necessity precludes taking the observation corresponding to the wave just before retirement to be included in the study, because then the age at retirement would be perfectly explained by the age one year before.

1986 interview, will not be included in the group of retirees because commuting information, which stopped to be collected in 1986, would be unknown for them. Instead, since only 18 subjects are in this situation, what we think renders useless the introduction of an additional statistical model for moves or job changes, they will be included in the group of attritors from the survey.

In this version of the paper, we follow standard practice and determine a study subject's labor force status from his answer to the question "Are you working now, unemployed, retired, or what?" We define respondents as being in the labor force if they identify themselves as currently working, unemployed/looking for work, or temporarily laid off. All other answers, including being permanently disabled, are considered indicative of retirement. The extension of the analysis to other notions of retirement (Lazear 1986) is left for future work. A study subject is considered an attritor if, before retiring, he leaves the survey or, as mentioned before, moves or changes job after the 1986 interview. Reasons for leaving the survey are that the entire family became non-response, that the subject died, or that a subject's move out of the family unit was not successfully followed. Regarding moves, we consider a change in residential location has taken place when answering "Yes" to "Have you head moved since last spring/previous interview?" Since the 1988 wave, a job change between interviews can be established using information on the beginning date with the present employer: "In what month and year did you start working for your present employer? (Count yourself as the employer if you are self-employed, and) give us your most recent start date if you have gone to work for them more than once". Before that wave, however, this precise information is not asked, and since the PSID generally provides no employer codes that uniquely identify jobs, establishing job changes is more complicated. We follow Brown and Light (1992) and assume that, before the 1988 wave, a job change has occurred whenever reported tenure is less than elapsed time since the previous interview.[14] In the subsequent analysis, attrition and retirement will be considered absorbing states.

In the survey waves collected between the data wave and the wave of retirement/attrition, most study subjects report commuting characteristics different than those reported in the data wave: 71.6 percent of subjects report different distances, and 83.3 percent report different times. While reporting errors may account for a proportion of the differences, the fact that the location of the workplace is unknown makes it possible that some fraction of the differences indicate genuine changes in distance or time to the workplace as a consequence, for example, of workers changing workplaces while doing the same job. Given the impossibility to include in the study the

[14] For the interview years 1969-1975, job tenure was coded in intervals. The interval code is translated into a precise value of tenure using the midpoint of the reported tenure interval.

observation provided just before retirement (see note 13), each study subject is assigned the average values of commuting distance and commuting time reported between the data wave and the wave of retirement/attrition. If reporting errors are random, they would be faded away with this procedure. Taking the average is also relevant if, when choosing a job, the worker knows he may have different commutes as a consequence of the job being developed in different places. Differences between commuting values reported in the data wave and those averaged thereafter are small: averages in the former case are 10.0 miles and 23.9 minutes, whereas averages of individual-level averages amount to 9.8 miles and 22.8 minutes.

3.2. Econometric model

The statistical model selected for analyzing the data has to be able to deal with non-random attrition, left-truncation, and endogenous regressors. The stochastic processes underlying labor market behavior and the behavior concerning participation in the PSID may be non-independent. This would happen if, for example, retirees move to Florida and are lost by the survey, for then workers attriting at a given wave would be at an unusually high risk of retiring and would therefore not be representative of those continuing in the survey. Second, as discussed before, each study subject contributes one observation to the study. Their age at that observation need not be the same, what raises a potential sample selection issue created by the different labor force participation rates across ages. Finally, the theoretical model underlined the possibility that residential location and labor supply were simultaneously decided. This, in turn, raises the possibility that interpersonal differences in commuting can be endogenous in a model for the age at retirement.[15]

We consider appropriate for analytical purposes the econometric model discussed in Wooldridge (2002, p. 567):

$$y_1 = \mathbf{z}_1\boldsymbol{\delta}_1 + \alpha_1 y_2 + u_1, \tag{6}$$

$$y_2 = \mathbf{z}\boldsymbol{\delta}_2 + u_2, \tag{7}$$

$$y_3 = 1(\mathbf{z}\boldsymbol{\delta}_3 + u_3 > 0).^{16} \tag{8}$$

In these expressions, y_1 represents (the log of) age at retirement, y_2 denotes commuting time or distance, y_3 is an indicator for non-attrition from the survey (before the moment of retirement), $\mathbf{z}$ is a vector of assumed exogenous covariates, $\mathbf{z}_1$ is a strict subset of $\mathbf{z}$;

[15] The sample design might also generate right-censored observations, that is, study subjects still participating in 2005. In this study, however, all subjects either retire or attrit before that date.

[16] A subindex denoting individuals is skipped for notational simplicity.

α_1, $\mathbf{\delta}_1$, $\mathbf{\delta}_2$, and $\mathbf{\delta}_3$ are parameters, whereas u_1, u_2, and u_3 represent arbitrarily correlated random errors. Equation (6), a linear regression for (the log of) age at retirement, is the equation of interest, equation (7) is a linear projection for the potentially endogenous commuting measure, and equation (8) is a selection equation. Variable y_1 is only observed when $y_3 = 1$, that is, when the study subject remains in the survey at least until withdrawing the labor force.

The equilibrium age-at-retirement gradient discussed in Section 2 concerned the location and retirement behavior of identical individuals. Hence, study subjects' heterogeneous characteristics are to be controlled for. Included in $\mathbf{z}_1$ are covariates emphasized by previous research on the timing of retirement (among many others, Mincer 1974, Filer and Petri 1988, Bound et al. 1996, and Coile 2004) and that could also affect location decisions: earnings, asset income, occupation, schooling, capacity for work, marital status, wife's employment status, and family size. Unfortunately, information on pension and social security wealth was not collected in the PSID during most of the survey waves included in this study. Even so, social security and pension benefit levels depend on previous earnings, and the more educated are more likely to have a pension (Taubman 1981). Moreover, $\mathbf{z}_1$ includes a set of year-of-birth dummies, which control, e.g., for changes in social security provisions that differently affected different cohorts of workers. Capacity for work is proxied by responses to the question: "Do you have a physical or nervous condition that limits the type of work, or the amount of work you can do?" Although there are a number of reasons to be suspicious about self-reported work limitations (Bound et al. 1995), alternative health measures are not available for most survey waves used in this study. Income variables are measured in real terms, with amounts deleted if they are below the corresponding first percentile or above the corresponding 99th percentile. Asset income is made up of rents, interests, dividends, and the asset part of business income; if a study subject is married or cohabiting, his income from assets includes the asset income of his partner. Potential left-truncation in retirement age is controlled for with a set of age-at-data-wave dummies included in $\mathbf{z}_1$. The difficulty to obtain data on price levels for other consumption goods and, specially, housing space (a commodity whose price varies across the urban space), make these variables to be excluded from the econometric model for age at retirement and Q and X to be assumed weakly separable from L.

Study subjects' observations pertain to different urban areas and to different time periods. On the one hand, as expression (5) suggests, the size of the age-at-retirement gradient might differ across urban areas if, for example, C', the derivative of the commuting expenses schedule, changes across areas (e.g., as a consequence of different urban transportation systems). On the other hand, the fact of having observations corresponding to different time periods opens the possibility that an urban area's level of commuting expenses for a given distance may have changed. The solution to the gradient heterogeneity problem is complicated. Survey confidentiality commitments preclude knowing study subjects' precise areas of residence, and even if these areas were known, the presence of just a few observations per area would preclude obtaining precise area-level estimates. Since the structure of transportation networks seems an important determinant of C', I will instead run a pooled estimation interacting commuting variables with the size of the largest city in the area of residence, assuming that central cities of similar size have similar transportation networks. A change in the level of commuting expenses for a given distance, $C(D)$, may be modeled as a time effect. Since by knowing the year of birth and the age at the data wave we may know the year the data wave belongs to, the sets of year-of-birth and age-at-data-wave dummies included in $\mathbf{z}_1$ do control for time effects.

The main explanatory variable in equation (6), y_2, represents either commuting distance or commuting time. Although both dimensions are highly connected (their sample correlation amounts to 0.75), they could differ in terms of amenability to policy intervention, rendering important to know the effects exerted by each dimension on the timing of retirement. As noted before, commuting characteristics can be endogenous in a model for age at retirement. I propose study subjects' race as an instrument for commuting. One of the main assumptions of the spatial mismatch hypothesis literature (see Kain 1992, Ihlanfeldt and Sjoquist 1998, and Gobillon et al. 2007 for surveys of this literature) is that blacks are disconnected from jobs: blacks mainly remained in the centres of U.S. urban areas in spite of suburbanization of jobs, especially entry-level jobs. Statistics strongly support this assertion (Gobillon et al. 2007). Two main explanations have been offered for understanding the causes of blacks' residential segregation and residential inertia in U.S. urban areas. The first, in line with the spatial mismatch hypothesis, emphasizes racial discrimination in the housing market, whereas the second emphasizes individuals' racial preferences leading to the spatial separation of ethnic groups. Both explanations seem unrelated to preferences for retirement, however, suggesting that race could be a valid instrument for commuting characteristics.[17] It is well-known, however, that the labor force participation rates of older, working-aged black men have historically been lower than those of white men,

[17] Another possible instrument would be previous commuting characteristics. For using these as instruments of current commuting, however, two conditions must be met: on the one hand, study subjects must have changed residence or job between the beginning of the study and the wave of retirement (in our sample, only 48% did); on the other hand, the error in a regression equation for expected age at retirement (like (6), but with expected age as dependent variable) should be uncorrelated across periods (information on expected retirement age was not collected in most PSID survey waves).

what suggests that blacks might have a preference for early withdrawal from the labor force. However, as shown by Bound et al. (1995, 1996), health differences between blacks and whites can account for most of the racial gap in labor force attachment for men. (Indeed, it is also possible that commuting differences be partly responsible for this racial gap.)

The estimation method is therefore as follows. We fist obtain $\hat{\boldsymbol{\delta}}_3$ from probit of y_3 on $\mathbf{z}$ using all observations (retirees and attritors). Although not necessary for identifying α_1 and $\boldsymbol{\delta}_1$, we shall include the (log of the) number of calls to secure an interview (Lillard and Panis 1998) as an exogenous variable affecting non-attrition. The inclusion of this variable in $\mathbf{z}$ allows assessing the validity of race as an instrument for commuting by performing a test of overidentifying restrictions. We then obtain the estimated inverse Mills ratios, $\hat{\lambda}_3 = \phi\left(\mathbf{z}\hat{\boldsymbol{\delta}}_3\right) \big/ \Phi\left(\mathbf{z}\hat{\boldsymbol{\delta}}_3\right)$, where $\phi(\cdot)$ and $\Phi(\cdot)$ represent the pdf and cdf of the standard normal distribution, respectively. Finally, using the subsample of retirees, we estimate the equation

$$y_1 = \mathbf{z}_1\boldsymbol{\delta}_1 + \alpha_1 y_2 + \gamma_1\hat{\lambda}_3 + error \tag{9}$$

by Two Stage Least Squares (2SLS) using instruments $\left(\mathbf{z}, \hat{\lambda}_3\right)$.

The final sample contains 1,045 study subjects, of whom 755 retire and 290 drop out from the survey before retirement. Sample selection criteria are listed in the data appendix. Table 1 offers a set of descriptive statistics. On average, retirement takes place in 1981, whereas, as a consequence of the sample design, attrition occurs on average earlier, in 1979. Between 1969 and the year of retirement or attrition, 35.5 percent of all subjects move, and near 29 percent change job at least once. On average, the data wave corresponds to the 1972 survey wave. Those who retire do so at about 62 years old, with time and distance to the workplace averaging 23 minutes and near 10 miles in the waves before retirement, both one-way figures. In total, 26.7 percent of the subjects are black.

3.3. Results

Tables 2 and 3 show the output of intermediate regressions: the probit for non-attrition before retirement (equation 8) and the first-stage regressions for the potentially endogenous commuting characteristics (equation 7), respectively. As in previous analyses of attrition in the PSID (e.g., Fitzgerald et al. 1998), the R-squared from the probit is small:[18] most effects are insignificant, and significant correlates of non-attrition

[18] This R-squared measure, defined in the notes to Table 2, is a usual measure of fit in binary-choice models.

explain little of attrition in the sample. Being married or black significantly reduces the probability of abandoning the survey, whereas the number of calls needed to secure an interview in the data wave has a strong correlation with future non-response.

Table 1. Descriptive statistics of data

	Retirees		Drop outs	
	Mean	Std dev	Mean	Std dev
Age at event (retirement/attrition)	61.7	5.1	56.7	6.6
Year of event (retirement/attrition)	1981	6	1979	7
Year of data collection (data wave)	1972	5	1972	4
Distance to workplace (miles, one-way)	9.9	10.5	9.5	11.1
Time to workplace (minutes, one-way)	23.0	18.1	22.3	18.4
Yearly earnings ($1,000)[a]	22.5	15.1	23.6	15.4
Yearly asset income ($1,000)[a]	1.4	3.5	1.5	4.1
Married (percent)	91.5	27.9	87.2	33.4
Wife not working (percent)	42.8	49.5	43.8	49.7
Family > 2 persons (percent)	62.5	48.4	68.3	46.6
Disabled (percent)	17.2	37.8	12.4	33.0
Less than high school education (percent)	55.4	49.7	56.6	49.7
More than high school education (percent)	11.4	31.8	14.1	34.9
Black (percent)	28.5	45.1	22.4	41.8
Number of calls to secure an interview	2.5	2.1	2.8	2.1
Moved between 1969 and year of event (percent)	35.5	47.9	35.5	47.9
Changed job between 1969 and year of event (%)	28.9	45.3	28.6	45.3
Living in area with largest city of (percent)				
500,000 or more inhabitants	31.8	46.6	33.8	47.4
50,000 – 499,999 inhabitants	33.1	47.1	35.2	47.8
Less than 50,000 inhabitants	35.1	47.6	31.0	46.3
Occupation (percent)				
Professional and technical workers	25.3	46.5	28.6	45.3
Sales and clerical workers	9.9	29.9	9.0	28.6
Craftsmen and operators	42.1	49.4	41.0	49.3
Service workers	18.8	39.1	16.9	37.5
Farm managers and laborers	3.8	19.2	4.5	20.7
Individuals	755		290	

Notes: [a] 1982-1984 dollars

Table 2. Probit for non-attrition (estimated coefficients)

Earnings ($1,000)	.0026 (.0039)
Asset income ($1,000)	-.0094 (.0132)
Married	.4960 (.1668)*
Wife not working	-.1414 (.0959)
Family > 2 persons	.0636 (.1141)
Disabled	.1825 (.1268)
Less than high school	-.1675 (.1099)
More than high school	-.1990 (.1698)
Sales and clerical worker	-.0524 (.1770)
Craftsman or operator	-.0288 (.1352)
Service worker	-.0660 (.1719)
Farm manager or laborer	-.2662 (.2508)
Black	.2669 (.1203)*
Log of number of calls	-.1597 (.0715)*
Intercept	-.3796 (.3000)
Observations	1,045
Drop outs	290
Log-likelihood	-554.90
R^2	.1009

Notes: Estimation includes birth-year dummies and age-at-data-wave dummies. Standard errors are in parentheses. R^2 equals one minus the ratio of the log likelihood of the fitted function to the log likelihood of a function with only an intercept. * Significant at 5%.

Table 3 shows the estimated linear projections for commuting distance (in its first column) and commuting time (second column). In line with previous findings on U.S. families' location patterns by income (e.g., Glaeser et al. 2008), we find that higher earnings are positively correlated with distance/time to the workplace. Income from assets, however, tends to be negatively correlated with time. As in Madden (1980), being married has a strong positive effect on commuting distance/time, though the wife's employment status is insignificant for commuting. The former effect could be due to married men locating in the suburbs as a consequence of having more children (housing space is cheaper there) and/or better financial prospects (assuming the income elasticity of housing demand exceeds the elasticity of commuting cost with income) than unmarried men. Since being more than two persons in the family unit appears unrelated to commuting, the second explanation seems more credible. The insignificant effect of wife's employment status on commuting distance/time could be the consequence of a pair of counteracting forces: on the one hand, a working wife would increase household income (though the difference in income between two-earner and traditional households is not big because the participation rate of married women is inversely related to husband's earnings (Hekman 1980)), generating a tendency to live further from the CBD if the income elasticity of housing demand exceeds the elasticity of commuting cost with income; on the other hand, two-earner households would be pull closer to the CBD to avoid doubling commuting costs. Interestingly, while having a limitation in the capacity to work is not correlated with commuting distance, it is associated with an increase in (one-way) commuting time of approximately 5 minutes. Having limitations in the kind and/or amount of work that can be done might indicate the presence of limitations in the ability to travel around, which could influence the choice of commuting mode and/or reduce the walking speed (thus, for example, increasing the walking time between the parking lot and the home/workplace). By occupations, farmers are by far those showing shorter commutes. The statistical significance of the inverse Mills ratio in the equation for commuting time suggests that unobserved characteristics affecting the decision to abandon the survey are related to unobserved characteristics affecting the duration of commutes. Most importantly for our purposes, we find that, other things constant, blacks have longer commutes than non-blacks (these made up essentially of whites): being black is associated to living about 3 miles further from the workplace and to devoting about 13 minutes more to a one-way commute. Other multivariate studies finding longer commutes for blacks include Chung et al. (2001) and Gabriel and Rosenthal (1996). The former study, using 1980 and 1990 Census Public Use Microdata Samples for Chicago and Los Angeles, finds that the commute times for Hispanics and whites were about the same, while blacks' was higher than either. Similarly, findings in Gabriel and Rosenthal (1996), who use data from the 1985 and 1989 American Housing Survey, indicate that, even after adjusting for neighbourhood fixed effects, earnings, and education, black workers have significantly longer commutes than comparable Asian and white workers. Table 3 also reports the value of the F statistic for the test that the instruments (being black as well as the number of calls to secure an interview) do not enter the first-stage regressions. Instruments are weakly correlated with distance, but much more correlated with time (it is indeed difficult to explain for what reason the number of calls to secure an interview is statistically significant—at the 10%—in the equation for commuting time).

Table 4 shows the output of estimating the model in equation (9) by OLS (in its first and second columns) and 2SLS (columns 3 and 4). The columns with the 2SLS estimates report also the value of the Sargan's statistic for testing the validity of the excluded instruments. This statistic is calculated as the number of observations times the R-squared from a regression of the 2SLS residuals on the full set of instruments. The null hypothesis that the instruments are valid cannot be rejected, with p-values well above standard significance levels. In general, OLS estimated coefficients tend to be rather similar to those estimated by 2SLS. We have indeed performed a formal test of endogeneity of commuting distance/time by extending equation (9) with residuals from regressions in columns (1) and (2) of Table 3 (one residual at a time), and then testing for the statistical significance of this extra variable. The null hypothesis that commuting characteristics are exogenous cannot be rejected at standard significance levels (p-values are 0.411 for commuting distance and 0.324 for commuting time). We therefore concentrate the discussion on the estimates shown in columns (1) and (2) of Table 4, as OLS is more efficient than 2SLS in the present context.

Table 3. Estimated linear projections for commuting characteristics

	Dependent variable:	
	(1)	(2)
	Distance (miles, one-way)	Time (minutes, one-way)
Earnings ($1,000)	.1001 (.0365)**	.2685 (.0605)**
Asset income ($1,000)	-.1265 (.1277)	-.3948 (.2114)*
Married	6.433 (3.066)**	17.218 (5.076)**
Wife not working	.7377 (1.121)	-1.201 (1.857)
Family > 2 persons	-.8875 (1.002)	.1130 (1.660)
Disabled	.4587 (1.371)	5.110 (2.270)**
Less than high school	-.1395 (1.314)	-2.637 (2.176)
More than high school	-2.185 (1.856)	-8.879 (3.072)**
Sales and clerical worker	-.0897 (1.542)	-1.038 (2.552)
Craftsman or operator	.9150 (1.199)	1.383 (1.984)
Service worker	2.505 (1.539)	4.597 (2.548)*
Farm manager or laborer	-5.660 (2.650)**	-19.795 (4.387)**
$\hat{\lambda}_3$	9.827 (11.909)	57.292 (19.717)**
Black	3.110 (1.706)*	13.250 (2.825)**
Log of number of calls	-.3116 (.9933)	-3.122 (1.645)*
Intercept	-43.72 (19.57)**	-5.416 (11.82)
F (excluded instruments)	2.37	12.30
R^2	0.115	0.184
Observations	755	

Notes: All estimations include birth-year dummies and age-at-data-wave dummies. Standard errors are in parentheses. * Significant at 10%. ** Significant at 5%.

Table 4. OLS and 2SLS regressions for age at retirement

	Dependent variable: log of age at retirement			
	(1)	(2)	(3)	(4)
	OLS	OLS	2SLS	2SLS
Miles, one-way	-.00048 (.00027)*		.0022 (.0033)	
Minutes, one-way		-.00021 (.00016)		.0006 (.0008)
Earnings ($1,000)	.00069 (.00024)**	.00069 (.00024)**	.00048 (.00036)	.00055 (.00027)**
Asset income ($1,000)	.0003 (.0009)	.0003 (.0009)	.0006 (.0009)	.0005 (.0009)
Married	.0156 (.0140)	.0144 (.0140)	.0072 (.0178)	.0131 (.0138)
Wife not working	-.0001 (.0062)	-.0002 (.0063)	-.0041 (.0082)	-.0026 (.0066)
Family > 2 persons	.0120 (.0069)*	.0123 (.0069)*	.0141 (.0076)*	.0124 (.0067)*
Disabled	-.0230 (.0077)**	-.0227 (.0077)**	-.0213 (.0082)**	-.0226 (.0076)**
Less than high school	-.0010 (.0070)	-.0008 (.0071)	-.0037 (.0080)	-.0033 (.0074)
More than high school	.0247 (.0110)**	.0246 (.0110)**	.0273 (.0118)**	.0269 (.0110)**
Sales or clerical worker	-.0156 (.0106)	-.0155 (.0106)	-.0168 (.0110)	-.0167 (.0105)
Craftsman or operator	-.0062 (.0083)	-.0062 (.0083)	-.0098 (.0096)	-.0088 (.0086)
Service worker	-.0080 (.0101)	-.0077 (.0101)	-.0178 (.0160)	-.0155 (.0126)
Farm manager or laborer	.0164 (.0166)	.0157 (.0167)	.0282 (.0225)	.0267 (.0196)
$\hat{\lambda}_3$	.0222 (.0412)	.0222 (.0412)	.0292 (.0432)	.0270 (.0407)
Intercept	4.008 (.0460)**	4.008 (.0461)**	3.988 (.053)**	3.993 (.0476)**
Sargan's statistic (1 over-identifying restriction)			.074 [.786]	.011 [.915]
R^2	.345	.343	.251	.317
Observations	755			

Notes: All estimations include birth-year dummies and age-at-data-wave dummies. Standard errors are in parentheses and probability values are in brackets. * Significant at 10%. ** Significant at 5%.

We find negative estimates associated to commuting variables. Although, as discussed before, a certain fraction of the sample commutes might not be trips to the CBD, but circumferential trips not ending in the CBD, this result suggests that the sign of expression (5) is positive. The estimated effects of commuting are small, being statistically different from zero at the 10% in the case of distance to the workplace only. As a consequence, the role of the age-at-retirement gradient in the urban space as an explanatory factor for earlier retirement is not significant: since a one-mile increase in the home-workplace separation is estimated to reduce age at retirement by approximately 0.05 percent, the increase in the home-workplace separation of about 6 miles observed among U.S. older workers in the last three decades of the 20th century would have caused a reduction of about 0.3% in retirement age (somewhat more than 2 months if initial retirement age were 65 years old). Keeping other things constant, higher earnings lead to later retirement, suggesting that the substitution effect of an earnings increase dominates the income effect. Later retirement ages are also found among male heads of families with more than two persons and among those more educated, this latter effect perhaps as a consequence of more educated individuals having steeper age-earnings profiles (Mincer 1974). Suffering a limitation in the capacity for work is estimated to reduce retirement age by a significant 2.3 percent. We find no evidence of non-random attrition from the PSID: the coefficient associated to the inverse Mills ratio is statistically not different from zero.

Finally, Table 5 offers the (selected) output of a pair of OLS estimations of equation (9) where the commuting variables (distance in column 1, time in column 2) have been interacted with dummies for the size of the biggest city in area. As discussed previously, the size of the age-at-retirement gradient might be heterogeneous across urban areas. To control in some extent for this possibility, three dummy variables representing areas whose biggest city contains 500,000 inhabitants or more (big city), between 50,000 and 499,999 inhabitants (medium city), and less than 50,000 inhabitants (small city) have been created. In the estimations, the interaction of commuting variables with big city is the unreported category. No significant differences in the size of the age-at-retirement gradient across urban areas of different size are detected.

Table 5. OLS regressions for age at retirement (selected estimated coefficients)

	Dependent variable: log of age at retirement	
	(1)	(2)
	OLS	OLS
Miles, one-way	-.00074 (.00039)*	
Miles*medium city in area	.00028 (.00051)	
Miles*small city in area	.00043 (.00044)	
Minutes, one-way		-.00028 (.00017)
Minutes*medium city in area		-.00001 (.00024)
Minutes*small city in area		.00029 (.00023)
R^2	.345	.345
Observations	755	

Notes: Estimations include controls reported in Table 4, birth-year dummies, and age-at-data-wave dummies. Standard errors are in parentheses. * significant at 10%

4. CONCLUSION

This paper has made two points. By extending a standard model of residential location and labor supply, we have shown that commuting and retirement can be causally related as a consequence of an equilibrium age-at-retirement gradient in the urban space: workers living further from the CBD, but being identical in all other respects to those residing closer, may retire at a different age as a consequence of the different commuting costs and housing space prices they face. It is an empirical matter to assess the sign and magnitude of this gradient, but if the commodity "years of leisure" is a normal good and if years of leisure and housing space are Slutsky complements (as wells as some additional conditions), workers living further from the CBD will retire earlier.

Secondly, we have quantified the magnitude of the age-at-retirement gradient for the U.S. economy, which allows assessing the role that the rise in the home-workplace separation observed among U.S. older male workers in the last three decades of the 20th century may have had in reducing their LFPR. A linear regression model for age at retirement has been estimated with data provided by the PSID and using an instrumental variables approach: commuting distance/time could be jointly decided with age at retirement, and is instrumented with workers' race. The validity of race as an instrument for commuting cannot be rejected, being race also sufficiently partially correlated (with commuting time) to be considered a reliable instrument. We have found some evidence of an effect of commuting on retirement, by which workers residing further from the workplace would tend to retire earlier, but the size of the effect seems insufficient for the increase in home-workplace separation to partly account for the decreasing trend in male LFPR in the U.S.

APPENDIX A. TECHNICAL APPENDIX

In this appendix we specify more fully the residential location and labor supply model outlined in Section 2, giving special emphasis to the mathematical derivation of the analytical results discussed in the main text. The worker's problem of choosing the most preferred residential location and retirement age given constraints can be formalized as a classical consumer demand problem. Let the worker's preferences be summarized by the utility function

$$U = U\left(L, Q, X, D\right), \tag{A1}$$

which is assumed twice continuously differentiable, increasing in the first three arguments, decreasing in the fourth argument, and strictly quasiconcave. Let the matrix Φ be given by

$$\Phi = \begin{pmatrix} 0 & U_1 & U_2 & U_3 & U_4 \\ U_1 & U_{11} & U_{12} & U_{13} & U_{14} \\ U_2 & U_{21} & U_{22} & U_{23} & U_{24} \\ U_3 & U_{31} & U_{32} & U_{33} & U_{34} \\ U_4 & U_{41} & U_{42} & U_{43} & U_{44} \end{pmatrix}, \tag{A2}$$

where subscripts denote partial derivatives: for example, $U_1(L,Q,X,D) = \partial U(L,Q,X,D)/\partial L$ and $U_{12}(L,Q,X,D) = \partial U(L,Q,X,D)/\partial L\,\partial Q$. By the strict quasiconcavity of $U(\cdot)$, $\left|{}_3\Phi_3\right| > 0$, $\left|{}_4\Phi_4\right| < 0$, and $\left|\Phi\right| > 0$, where ${}_r\Phi_r$ denotes the $r \times r$ submatrix of Φ where only the first r rows and r columns are retained. In maximizing utility, workers are subject to the time constraint

$$T = L + H \tag{A3}$$

and the budget constraint

$$X + R\left(D\right)Q + C\left(D\right)H \le WH + Y, \tag{A4}$$

which can be combined in the form of a full-income budget constraint:

$$X + R\left(D\right)Q + \left(W - C\left(D\right)\right)L \le F \equiv \left(W - C\left(D\right)\right)T + Y. \tag{A5}$$

A1. Equilibrium

In order to facilitate the discussion of the equilibrium, the worker's decision-making process is decomposed into two stages. In the first stage, optimal quantities of L, Q, and X are chosen for a given distance, obtaining a constrained or short-run maximum of utility. In the second stage, distance itself is chosen so as to maximize utility.

At a distance D from the CBD, workers solve

$$\max_{L,Q,X} \; U = U\left(L, Q, X; D\right) \tag{A6}$$

$$\text{s. t. } X + RQ + \left(W - C\right)L \le F \equiv \left(W - C\right)T + Y \,, \tag{A7}$$

where dependence of R and C on D has been dropped for notational simplicity. The Lagrangean function of this optimization model is

$$U\left(L, Q, X; D\right) + \lambda\left(\left(W - C\right)T + Y - X - RQ - \left(W - C\right)L\right), \tag{A8}$$

and the necessary first-order conditions are

$$\left(W - C\right)T + Y - X - RQ - \left(W - C\right)L = 0 \,, \tag{A9}$$

$$U_1 - \lambda\left(W - C\right) = 0 \,, \tag{A10}$$

$$U_2 - \lambda R = 0 \,, \tag{A11}$$

$$U_3 - \lambda = 0 \,. \tag{A12}$$

From (A10)-(A12), the usual equilibrium conditions stating that the marginal rate of substitution between two goods equals the ratio of their marginal costs are obtained. Let the matrix P be

$$\mathrm{P} = \begin{pmatrix} 0 & \left(W - C\right) & R & 1 \\ \left(W - C\right) & U_{11} & U_{12} & U_{13} \\ R & U_{21} & U_{22} & U_{23} \\ 1 & U_{31} & U_{32} & U_{33} \end{pmatrix}, \tag{A13}$$

where second-order partial derivatives are evaluated at some point satisfying the first-order necessary conditions. Sufficient conditions for a constrained optimum are $\left|{}_3\mathrm{P}_3\right| > 0$ and $\left|\mathrm{P}\right| < 0$. Using expressions (A10)-(A12), we see that $\left|{}_3\mathrm{P}_3\right| = \dfrac{1}{\lambda^2}\left|{}_3\Phi_3\right|$ and $\left|\mathrm{P}\right| = \dfrac{1}{\lambda^2}\left|{}_4\Phi_4\right|$. Hence, the strict quasiconcavity of $U\left(\cdot\right)$ guarantees that the sufficient conditions for an optimum are met. First-order conditions (A9)-(A12) have continuous partial derivatives with respect to all endogenous and exogenous variables, and the following Jacobian determinant

$$J = \begin{vmatrix} 0 & -\left(W - C\right) & -R & -1 \\ -\left(W - C\right) & U_{11} & U_{12} & U_{13} \\ -R & U_{21} & U_{22} & U_{23} \\ -1 & U_{31} & U_{32} & U_{33} \end{vmatrix} \tag{A14}$$

evaluated at any values of the variables is nonzero: $J = \left(-1\right)^2 \left|\mathrm{P}\right| = \left(-1\right)^2 \dfrac{1}{\lambda^2}\left|{}_4\Phi_4\right| \neq 0$. Then, by the implicit function theorem, for some neighborhood of the solution there exist implicit short-run demand functions

$$L = L\left(F,(W-C),R;D\right), \tag{A15}$$

$$Q = Q\left(F,(W-C),R;D\right), \tag{A16}$$

$$X = X\left(F,(W-C),R;D\right), \tag{A17}$$

generally expressed as

$$Z_j = Z_j\left(F,P_1,P_2;D\right), \tag{A18}$$

with $Z_1 = L$, $Z_2 = Q$, $Z_3 = X$, $P_1 = (W-C)$, and $P_2 = R$.

The constrained maximum of utility obtained at a distance D from the CBD is denoted $\bar{U}(D) = U\left(L,Q,X;D\right)$, where L, Q, and X are the utility-maximizing demands satisfying expressions (A9)-(A12). In order to know how changes in D alter the constrained optimum, we firstly use the chain rule:

$$\frac{\partial \bar{U}(D)}{\partial D} = U_1 \frac{\partial L(D)}{\partial D} + U_2 \frac{\partial Q(D)}{\partial D} + U_3 \frac{\partial X(D)}{\partial D} + U_4. \tag{A19}$$

Then, by the envelope theorem,

$$\frac{\partial \bar{U}(D)}{\partial D} = -\lambda\left(QR' + HC'\right) + U_4. \tag{A20}$$

A full equilibrium is thus attained when expression (A20) equals zero, meaning that workers cannot increase utility by changing their residential location. Jointly with the first-order conditions for a constrained optimum, equation (A20) forms a system of five equations in five unknowns that can be solved for the utility-maximizing demand equations and an equation determining the optimum residential location:

$$Z_j = \Psi_j\left(F,P_1,P_2\right) \tag{A21}$$

$$D = \Psi_D\left(F,P_1,P_2\right). \tag{A22}$$

Demand equations (A21) represent long-run demands where the worker has previously adjusted its location with (A22) to maximize utility.

Rewriting the equilibrium condition for the second step of utility maximization, we obtain the expression shown in the main text:

$$-QR' = -\frac{U_4}{\lambda} + HC'. \tag{A23}$$

Before discussing the comparative statics of the model, the direction of the worker's move in case both sides of (A23) are of different size is to be determined. We assume that

$$-QR'' \leq -\frac{U_{44}}{\lambda} + HC'', \tag{A24}$$

i.e., the slope of the marginal commuting costs curve is no lower than the slope of the marginal housing savings curve. Then, if for example $-QR' < -\frac{U_4}{\lambda} + HC'$, workers must move closer to the CBD to restore equilibrium. As argued by Muth (1969, p.25), condition (A24) makes the optimum distance to be spatially stable.

A2. Comparative-static analysis

In deriving the comparative-static results of the model, we work towards a twofold goal. On the one hand, as usual, we assess the consequences for endogenous variables of changes in model parameters. On the other hand, we assess the consequences for endogenous variables of changes in the equilibrium distance. As distance to the CBD is under the worker's control, it might seem we are mistakenly labeling this part of the analysis as comparative-static. Yet, given the structure of the model, a change in distance is formally equivalent to an exogenous change in the price of housing and the level of commuting costs, for example. Thus, it may be properly considered here.

Long-run responses to variations in parameters can be related to constrained, short-run responses by means of the chain rule. For example:

$$\left(\frac{\partial Z_j}{\partial P_i}\right)_{LR} = \left(\frac{\partial Z_j}{\partial P_i}\right)_D + \left(\frac{\partial Z_j}{\partial D}\right)\left(\frac{\partial D}{\partial P_i}\right). \tag{A25}$$

The total effect of a change in a parameter may be conceptually decomposed into two effects. On the one hand, the short-run effect $\left(\left(\frac{\partial Z_j}{\partial P_i}\right)_D\right)$. On the other hand, the effect caused by the variation in housing prices and commuting costs as a consequence of the move induced by the change in the parameter $\left(\left(\frac{\partial Z_j}{\partial D}\right)\left(\frac{\partial D}{\partial P_i}\right)\right)$. We are particularly interested in the responses $\left(\frac{\partial L}{\partial D}\right)$ and $\left(\frac{\partial L}{\partial C}\right)_{LR}$. The next paragraphs lay out the pieces that allow assessing the effects of changes in parameters on long-run demands.

The properties of the constrained demand functions in (A15)-(A17) can be expressed in terms of full-income, μ_j, and partial substitution effects, s_{ji}:

$$\mu_j = \left(\frac{\partial Z_j}{\partial F}\right), \qquad s_{ji} = \left(\frac{\partial Z_j}{\partial P_i}\right)_{\bar{U}}, \tag{A26}$$

where the subscript $\bar{U}$ indicates a compensated price response holding (constrained) utility constant. Totally differentiating the first-order conditions (A9)-(A12) we have

$$\begin{bmatrix} 0 & -(W-C) & -R & -1 \\ -(W-C) & U_{11} & U_{12} & U_{13} \\ -R & U_{21} & U_{22} & U_{23} \\ -1 & U_{31} & U_{32} & U_{33} \end{bmatrix} \begin{bmatrix} d\lambda \\ dL \\ dQ \\ dX \end{bmatrix} = \begin{bmatrix} -1 & -H & Q & H \\ 0 & \lambda & 0 & -\lambda \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} dY \\ dW \\ dR \\ dC \end{bmatrix}. \tag{A27}$$

Assuming that only non-labor income varies:

$$\begin{bmatrix} 0 & -(W-C) & -R & -1 \\ -(W-C) & U_{11} & U_{12} & U_{13} \\ -R & U_{21} & U_{22} & U_{23} \\ -1 & U_{31} & U_{32} & U_{33} \end{bmatrix} \begin{bmatrix} \partial\lambda/\partial Y \\ \partial L/\partial Y \\ \partial Q/\partial Y \\ \partial X/\partial Y \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{A28}$$

By Cramer's rule,

$$\frac{\partial L}{\partial Y} = \frac{1}{J} \begin{vmatrix} 0 & -1 & -R & -1 \\ -(W-C) & 0 & U_{12} & U_{13} \\ -R & 0 & U_{22} & U_{23} \\ -1 & 0 & U_{32} & U_{33} \end{vmatrix} = \frac{(-1)^2 \frac{1}{\lambda^2} \lambda \left({}_4\Phi_4^{12}\right)}{\frac{1}{\lambda^2} \left|{}_4\Phi_4\right|} = \frac{\lambda\left({}_4\Phi_4^{12}\right)}{\left|{}_4\Phi_4\right|}, \tag{A29}$$

where ${}_4\Phi_4^{12}$ is the adjoint of element $(1,2)$ in ${}_4\Phi_4$. Also,

$$\frac{\partial Q}{\partial Y} = \frac{\lambda\left({}_4\Phi_4^{13}\right)}{\left|{}_4\Phi_4\right|}, \qquad \frac{\partial X}{\partial Y} = \frac{\lambda\left({}_4\Phi_4^{14}\right)}{\left|{}_4\Phi_4\right|}. \tag{A30}$$

Non-labor income effects coincide with full-income effects, for $\frac{\partial Z_j}{\partial Y} = \frac{\partial Z_j}{\partial F}\frac{\partial F}{\partial Y} = \frac{\partial Z_j}{\partial F}$.

Assuming that only the wage rate varies,

$$\begin{bmatrix} 0 & -(W-C) & -R & -1 \\ -(W-C) & U_{11} & U_{12} & U_{13} \\ -R & U_{21} & U_{22} & U_{23} \\ -1 & U_{31} & U_{32} & U_{33} \end{bmatrix} \begin{bmatrix} \partial\lambda/\partial W \\ \partial L/\partial W \\ \partial Q/\partial W \\ \partial X/\partial W \end{bmatrix} = \begin{bmatrix} -H \\ \lambda \\ 0 \\ 0 \end{bmatrix}. \tag{A31}$$

By Cramer's rule,

$$\frac{\partial L}{\partial W} = H\frac{\lambda\left({}_4\Phi_4^{12}\right)}{\left|{}_4\Phi_4\right|} + \frac{\lambda\left({}_4\Phi_4^{22}\right)}{\left|{}_4\Phi_4\right|} = H\mu_1 + s_{11}, \tag{A32}$$

$$\frac{\partial Q}{\partial W} = H\frac{\lambda\left({}_4\Phi_4^{13}\right)}{\left|{}_4\Phi_4\right|} + \frac{\lambda\left({}_4\Phi_4^{23}\right)}{\left|{}_4\Phi_4\right|} = H\mu_2 + s_{21}, \tag{A33}$$

$$\frac{\partial X}{\partial W}=H\frac{\lambda\left({}_4\Phi_4^{14}\right)}{\left|{}_4\Phi_4\right|}+\frac{\lambda\left({}_4\Phi_4^{24}\right)}{\left|{}_4\Phi_4\right|}=H\mu_3+s_{31}. \tag{A34}$$

Terms s_{j1} are partial substitution effects, for $\frac{\partial Z_j}{\partial W}=\frac{\partial Z_j}{\partial(W-C)}\frac{\partial(W-C)}{\partial W}=\frac{\partial Z_j}{\partial(W-C)}$.

Since $\frac{\partial Z_j}{\partial C}=\frac{\partial Z_j}{\partial(W-C)}\frac{\partial(W-C)}{\partial C}=-\frac{\partial Z_j}{\partial(W-C)}$, partial effects with respect to the level of commuting expenses are equal in magnitude, but opposite in sign, than those generated by changes in W.

Assuming that only R varies,

$$\begin{bmatrix} 0 & -(W-C) & -R & -1 \\ -(W-C) & U_{11} & U_{12} & U_{13} \\ -R & U_{21} & U_{22} & U_{23} \\ -1 & U_{31} & U_{32} & U_{33} \end{bmatrix}\begin{bmatrix} \partial\lambda/\partial R \\ \partial L/\partial R \\ \partial Q/\partial R \\ \partial X/\partial R \end{bmatrix}=\begin{bmatrix} Q \\ 0 \\ \lambda \\ 0 \end{bmatrix}. \tag{A35}$$

Again by Cramer's rule,

$$\frac{\partial L}{\partial R}=-Q\frac{\lambda\left({}_4\Phi_4^{12}\right)}{\left|{}_4\Phi_4\right|}+\frac{\lambda\left({}_4\Phi_4^{32}\right)}{\left|{}_4\Phi_4\right|}=-Q\mu_1+s_{12}, \tag{A36}$$

$$\frac{\partial Q}{\partial R}=-Q\frac{\lambda\left({}_4\Phi_4^{13}\right)}{\left|{}_4\Phi_4\right|}+\frac{\lambda\left({}_4\Phi_4^{33}\right)}{\left|{}_4\Phi_4\right|}=-Q\mu_2+s_{22}, \tag{A37}$$

$$\frac{\partial X}{\partial R}=-Q\frac{\lambda\left({}_4\Phi_4^{14}\right)}{\left|{}_4\Phi_4\right|}+\frac{\lambda\left({}_4\Phi_4^{34}\right)}{\left|{}_4\Phi_4\right|}=-Q\mu_3+s_{32}. \tag{A38}$$

In deriving the sign of $\left(\frac{\partial Z_j}{\partial D}\right)$, distance is treated as if it were another parameter, so that the procedure is formally identical to that just followed to obtain constrained demand responses. Short-run demands are totally differentiated with respect to D. If $U_{j4}=0,\ j=1,2,3$, an extra column with entries

$$\begin{bmatrix} QR'+HC' \\ -\lambda C' \\ \lambda R' \\ 0 \end{bmatrix} \tag{A39}$$

is added to the right-hand side matrix in (A27). Also, the term dD is to be added to the right-hand side vector in (A27) occupying a position in accordance to the position occupied by the column added to the matrix. Assuming that only D varies,

$$\begin{bmatrix} 0 & -(W-C) & -R & -1 \\ -(W-C) & U_{11} & U_{12} & U_{13} \\ -R & U_{21} & U_{22} & U_{23} \\ -1 & U_{31} & U_{32} & U_{33} \end{bmatrix} \begin{bmatrix} \partial\lambda/\partial D \\ \partial L/\partial D \\ \partial Q/\partial D \\ \partial X/\partial D \end{bmatrix} = \begin{bmatrix} QR' + HC' \\ -\lambda C' \\ \lambda R' \\ 0 \end{bmatrix}. \tag{A40}$$

By Cramer's rule,

$$\frac{\partial L}{\partial D} = -(QR' + HC')\mu_1 - C's_{11} + R's_{12}, \tag{A41}$$

$$\frac{\partial Q}{\partial D} = -(QR' + HC')\mu_2 - C's_{21} + R's_{22}, \tag{A42}$$

$$\frac{\partial X}{\partial D} = -(QR' + HC')\mu_3 - C's_{31} + R's_{32}. \tag{A43}$$

Finally, the sign of $\left(\frac{\partial D}{\partial P_i}\right)$ can be inferred from the sign of $\left(\frac{\partial \gamma}{\partial P_i}\right)_D$, where $\gamma \equiv -\frac{U_4}{\lambda} + HC' + QR'$. Variable γ stands for the net costs (respectively, net savings) of a very short move from the equilibrium location and away from (towards) the CBD. For example, starting from a position of equilibrium, an increase in γ due to some variation in a parameter would imply that, at the former equilibrium location, marginal commuting costs would now exceed marginal housing space savings, and, under condition (A24), the worker would move closer to the CDB to restore equilibrium. Assuming again $U_{j4} = 0,\ j = 1,2,3$, we have

$$\left(\frac{\partial \gamma}{\partial Y}\right)_D = C'\left(\frac{\partial H}{\partial Y}\right)_D + R'\left(\frac{\partial Q}{\partial Y}\right)_D + \frac{U_4}{\lambda^2}\left(\frac{\partial \lambda}{\partial Y}\right)_D. \tag{A44}$$

The first two terms on the right-hand are negative whenever a rise in non-labor income expands short-run demands for housing space and leisure. The sign of the last term, however, is not known (it would be positive if the utility function is strictly concave, for then $\left(\frac{\partial \lambda}{\partial Y}\right)_D < 0$). With regard to variations in prices,

$$\left(\frac{\partial \gamma}{\partial P_j}\right)_D = C'\left(\frac{\partial H}{\partial P_j}\right)_D + R'\left(\frac{\partial Q}{\partial P_j}\right)_D + \frac{U_4}{\lambda^2}\left(\frac{\partial \lambda}{\partial P_j}\right)_D. \tag{A45}$$

APPENDIX B. DATA APPENDIX

This appendix lists the sample selection criteria observed in the construction of the sample object of study. For purposes of replication, either marginal loses of observations or surviving number of observations are provided after applying each selection criterion.

The sample is obtained from a PSID data set made up of "All Individuals Data" files that were successively downloaded from the PSID Data Center (http://simba.isr.umich.edu/) between October 2006 and July 2008. This data set contains information for the 53,005 individuals interviewed by the PSID between 1968 and 1993, although included variables cover the period 1968-2005 (waves 1-34), and a total of 1,802,170 observations (person-years) are present in the data set. I keep only observations corresponding to heads or wives of PSID families (1,434,387 person-years lost). Then, I select men only, aged 41-60 in 1969, and that in 1969 are participating in the labor force. 1,134 persons satisfy these criteria, of whom 804 retire and 330 abandon the sample before retiring (including, as explained in the main text, 18 persons that move or change job after the 1986 interview). To be included in the final sample, these persons have to provide valid data in at least one survey wave since the data wave through the wave previous to retirement/attrition. (Invalid answering codes for commuting are 99 (in the case of miles) and 999 (in the case of time). Earnings and asset income below the corresponding 1st percentile or above the corresponding 99th percentile are deleted.) As a consequence of this criterion, 89 subjects are dropped and 1,045 remain in the sample, of whom 755 retire and 290 abandon the survey before retiring.

REFERENCES

Alonso, William. 1964. *Location and land use*. Cambridge, MA: Harvard University Press.

Ball, Michael J. 1973. Recent empirical work on the determinants of relative house prices. *Urban Studies* 10:213-233.

Blau, David, and Ryan Goodstein. 2007. What explains trends in labor force participation of older men in the United States? IZA Discussion Paper No. 2991.

Boskin, Michael J. 1977. Social security and retirement decisions. *Economic Inquiry* 15:1-25.

Bound, John, Michael Schoenbaum, and Timothy Waidmann. 1995. Race and education differences in disability status and labor force attachment in the Health and Retirement Survey. *The Journal of Human Resources* XXX:S227-S267.

Bound, John, Michael Schoenbaum, and Timothy Waidmann. 1996. Race differences in labor force attachment and disability status. *The Gerontologist* 36(3): 311-321.

Brown, James N., and Audrey Light. 1992. Interpreting panel data on job tenure. *Journal of Labor Economics* 10(3): 219-257.

Chung, Chanjin, Samuel L. Myers, Jr., and Lisa Saunders. 2001. Racial differences in transportation access to employment in Chicago and Los Angeles, 1980 and 1990. *The American Economic Review* 91(2): 174-177.

Cogan, John F. 1981. Fixed costs and labor supply. *Econometrica* 49:945-963.

Coile, Courtney C. 2004. Retirement incentives and couples' retirement decisions. *Topics in Economics Analysis & Policy* 4(1): Article 17.

Crampton, Graham R. 1999. Urban labour markets. In *Handbook of regional and urban economics*, ed. Paul Cheshire and Edwin S. Mills: North-Holland.

Filer, Randall K., and Peter A. Petri. 1988. A job-characteristics theory of retirement. *The Review of Economics and Statistics* 70:123-128.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998. An analysis of sample attrition in panel data. *The Journal of Human Resources* XXXIII(2): 252-299.

French, Eric. 2005. The effects of health, wealth, and wages on labour supply and retirement behaviour. *Review of Economic Studies* 72:395-427.

Gabriel, Stuart A., and Stuart S. Rosenthal. 1996. Commutes, neighborhood effects, and earnings: an analysis of racial discrimination and compensating differentials. *Journal of Urban Economics* 40:61-83.

Glaeser, Edward L., Matthew E. Kahn, and Jordan Rappaport. 2008. Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics* 63(1): 1-24.

Gobillon, Laurent, Harris Selod, and Yves Zenou. 2007. The mechanisms of spatial mismatch. *Urban Studies* 44(12): 2401-2428.

Gruber, Jonathan, and David A. Wise. 1999. *Social security and retirement around the world*. Chicago and London: The University of Chicago Press.

Gruber, Jonathan, and David A. Wise. 2004. *Social security programs and retirement around the world – micro-estimation*. Chicago and London: The University of Chicago Press.

Gustman, Alan L., and Thomas L. Steinmeier. 2000. Retirement in dual-career families: a structural model. *Journal of Labor Economics* 18:503-545.

Hekman, John S. 1980. Income, labor supply, and urban residence. 1980. *The American Economic Review* 70(4): 805-811.

Ihlanfeldt, Keith R., and David L. Sjoquist. 1998. The spatial mismatch hypothesis: a review of recent studies and their implications for welfare reform. *Housing Policy Debate* 9(4): 849-892.

Kahneman, Daniel, Alan B. Krueger, David A. Schkade, Norbert Schwarz, and Arthur A. Stone. 2004. A survey method for characterizing daily life experience: the day reconstruction method. *Science* 306:1776-1780.

Kain, John F. 1992. The spatial mismatch hypothesis: three decades later. *Housing Policy Debate* 3(2): 371-392.

Kolodziejczyk, Christophe. 2006. Retirement and fixed costs to work: an empirical analysis. Working Paper no. 2006-09, Centre for Applied Microeconometrics, University of Copenhagen.

Koslowsky, Meni, Avraham N. Kluger, and Mordechai Reich. 1995. *Commuting stress. Causes, effects, and methods of coping*. New York and London: Plenum Press.

Lazear, Edward P. 1986. Retirement from the labor force. In *Handbook of Labor Economics*, ed. Orley C. Ashenfelter and Richard Layard: Elsevier Science Publishers B.V.

Lillard, Lee A., and Constantijn W. A. Panis. 1998. Panel attrition from the Panel Study of Income Dynamics. *The Journal of Human Resources* XXXIII(2): 437-457.

Lumsdaine, Robin L., and Olivia S. Mitchell. 1999. New developments in the economic analysis of retirement. In *Handbook of Labor Economics*, ed. Orley C. Ashenfelter and David Card: North-Holland.

Madden, Janice Fanning. 1980. Urban land use and the growth in two-earner households. *The American Economic Review* 70(2): 191-197.

Madrian, Brigitte C., Gary Burtless, and Jonathan Gruber. 1994. The effect of health insurance on retirement. *Brookings Papers on Economic Activity* 1994(1): 181-252.

Mincer, Jacob. 1974. *Schooling, experience, and earnings*. New York: NBER.

Mitchell, Olivia M., and Gary M. Fields. 1984. The economics of retirement behavior. *Journal of Labor Economics* 2:84-105.

Muth, Richard F. 1969. *Cities and housing. The spatial pattern of urban residential land use*. Chicago and London: The University of Chicago Press.

Oi, Walter. 1976. Residential location and labor supply. *The Journal of Political Economy* 84, no. 4, pt. 2:S221-S238.

Rosen, Sherwin. 1986. The theory of equalizing differences. In *Handbook of Labor Economics*, ed. Orley C. Ashenfelter and Richard Layard: North-Holland.

Rust, John, and Christopher Phelan. 1997. How social security and medicare affect retirement behavior in a world of incomplete markets. *Econometrica* 65:781-831.

Schafer, Andreas. 2000. Regularities in travel demand: an international perspective. *Journal of Transportation and Statistics* 3(3):1-32.

Sheppard, Stephen. 1999. Hedonic analysis of housing markets. In *Handbook of regional and urban economics*, ed. Paul Cheshire and Edwin S. Mills: North-Holland.

Simpson, Wayne, and Anne van der Veen. 1992. The economics of commuting and the urban labour market. *Journal of Economic Surveys* 6:45-62.

Stock, James H., and David A. Wise. 1990. Pensions, the option value of work and retirement. *Econometrica* 58:1151-1180.

Straszheim, Mahlon. 1987. The theory of urban residential location. In *Handbook of regional and urban economics*, ed. Edwin S. Mills: North-Holland.

Stucker, James P. 1975. Transport improvements, commuting costs, and residential location. *Journal of Urban Economics* 2:123-143.

Taubman, Paul. 1981. Pensions and mortality, NBER Working Paper No. 811.

Wales, Terence J. 1978. Labour supply and commuting time. *Journal of Econometrics* 8:215-226.

White, Michelle J. 1999. Urban areas with decentralized employment: theory and empirical work. In *Handbook of regional and urban economics*, ed. Paul Cheshire and Edwin S. Mills: North-Holland.

Wooldridge, Jeffrey M. 2002. *Econometric analysis of cross section and panel data*. Cambridge and London: The MIT Press.